

Could Exposure Assessment Problems Give Us Wrong Answers to Nutrition and Cancer Questions?

Arthur Schatzkin, Victor Kipnis

Whether fruits and vegetables provide protection from cancer is of considerable public health importance—but it remains an open question. In this issue of the *Journal*, Hung et al. report findings from two cohort studies: fruit and vegetable intake is associated with a “modest” reduction in risk of noncommunicable (“chronic”) disease (1). This reduction is confined to cardiovascular disease. The association for cancer is null.

Most evidence for a protective role of fruits and vegetables against cancer has come from case-control studies. The authors rightly note the possibility of recall bias in such studies. Indeed, prospective cohort studies have tended to demonstrate weak or no associations between fruit and vegetable intake and malignant disease, including large-bowel cancer (2).

So the question arises: does total fruit and vegetable intake really confer little or no cancer protection? (In this editorial we address only the relation of total fruit and vegetable intake to all cancer. However, similar issues can be raised for associations between individual fruits or vegetables and cancers at specific sites.) Hung et al. consider—and largely dismiss—the possibility that problems with exposure assessment have caused their study to give the wrong answer to this question. However, more consideration of this possibility is in order.

The exposure assessment tool—food-frequency questionnaire (FFQ)—that is used to measure diet, including fruit and vegetable intake, is subject to substantial error, both random and systematic (3,4). Is this error sufficient to obscure an existing fruit and vegetable-cancer association? That is, could a true relative risk (RR), comparable to that seen by Hung et al. for cardiovascular disease (0.88 per increment of five daily servings of fruits and vegetables) be attenuated to the null by this error? Hung et al. argue against this possibility on three counts: First, they say that, by taking the cumulative average of several FFQs, they have reduced intraindividual random variation. This may be true, but such cumulative averaging doesn’t eliminate systematic error. Second, they assert that the FFQ has been validated against multiple weighted 1-week dietary records. We have argued elsewhere that such records are also subject to measurement error, and this error may well be correlated with that in the FFQ (4,5). Therefore the use of records as a reference instrument likely overestimates the accuracy of the food frequency questionnaire. Third, the authors assert that the observed association for cardiovascular disease means that the instrument they used is accurate enough to pick up any real association between fruit and vegetable intake and cancer. But suppose the true relative risk for cardiovascular disease is 0.75; attenuating 0.75 to 0.88 (for fruits and vegetables versus cardiovascular disease) is roughly comparable to attenuating 0.88 to 1.0 (for fruits and vegetables versus cancer).

And this is considering only the univariate problem, relating disease to a single variable (fruit and vegetable intake) measured

with error. The reality of nutritional epidemiologic research is multivariate: a dietary exposure such as total fruit and vegetable intake is generally examined in conjunction with energy intake and a variety of other covariates. The Hung et al. multivariate models include total energy intake, alcohol, smoking, vitamin supplement use, physical activity—13 variables in all for men and 16 for women. Measurement error affects determination of not only fruit and vegetable intake but also other dietary variables, such as alcohol and energy intake, and other covariates, such as physical activity. Moreover, the errors in measurement of these variables are likely to be correlated. The actual effect of this multivariate measurement error on true relative risk is complex. A true null relative risk could appear as an increased or decreased risk; a true protective association could be attenuated to the null (6). It would not take very much error in a few covariates (with perhaps a little intercorrelation of those errors) to attenuate a true protective association (RR = 0.88, for example) to the null.

Multivariate modeling in the presence of measurement error is not just an arcane statistical matter but a real concern in this field—especially if a serious attempt is being made to detect modest relative risks. Hung et al. observed a statistically significant relative risk of 0.92 for fruits and vegetables in a model including only age as a covariate. This relative risk was estimated as 0.97 (becoming statistically nonsignificant) when smoking was added to the model. They claim that the nearly null results in this simple model means that overadjustment by confounders is an unlikely explanation for the null findings in their multivariate model. However, even assuming (unrealistically) that measures of total fruit and vegetable intake contain only random error and that smoking is measured perfectly, adding smoking to the model would necessarily increase the attenuation due to random error in measurement of fruit and vegetable intake (7). This fact alone could potentially explain the additional attenuation of relative risk from 0.92 to 0.97. With more complex but more realistic measurement error, the situation could be even less predictable—and certainly consistent with attenuation of an important protective association.

Affiliations of authors: Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville, MD (AS); Biometry Research Group, Division of Cancer Prevention, National Cancer Institute, Bethesda, MD (VK).

Correspondence to: Arthur Schatzkin, MD, DrPH, Nutritional Epidemiology Branch, Division of Cancer Epidemiology and Genetics, National Cancer Institute, 6120 Executive Blvd., Ste. 320, Rm. 3040, Rockville, MD 20852-7232 (e-mail: schatzka@mail.nih.gov).

DOI: 10.1093/jnci/djh329

Journal of the National Cancer Institute, Vol. 96, No. 21, © Oxford University Press 2004, all rights reserved.

The effect of multivariate measurement error on estimated relative risk, moreover, depends not only on the size of and relations among measurement errors but also on the magnitude of true associations between disease and both the main exposure and covariates (7). These associations could differ for cardiovascular disease and cancer and thereby explain why, even if fruit and vegetable intake truly protected against both cardiovascular disease and cancer, an inverse association might be observed only for cardiovascular disease.

Another dietary assessment–related point: Can we be certain that using cumulative averaging of repeated dietary assessments is the optimum approach for capturing diet–disease relationships? Suppose more distant, long-term (and prestudy) diet is most relevant to carcinogenesis. Arguably, it is the baseline assessment that best reflects this prestudy diet. Using the cumulative average of repeated dietary assessments could introduce additional systematic measurement error and further distort diet–cancer relations.

What can researchers do about these potential exposure assessment problems?

First, appreciate them more. This means elevating the level of uncertainty that is attached to observed null or weak associations emerging in the multivariate context.

Second, routinely present results from sensitivity analyses based on both parsimonious and “kitchen-sink” multivariate models as well as alternative methods of handling multiple dietary assessments.

Third, improve measurement of dietary and nondietary variables. This includes obtaining better information on earlier life nutritional exposures (8) and possibly the use of instruments (e.g., food diaries or internet-based dietary recalls) other than FFQs for individual-level dietary assessments (9,10).

Fourth, incorporate intake biomarker–based calibration studies into prospective cohort studies to evaluate and possibly adjust for the effects of measurement error on observed diet–cancer associations (11).

Fifth, investigate the measurement error characteristics of dietary patterns and indexes (12,13,14). Use of such multifactorial approaches may reduce exposure misclassification, although this hypothesis is wholly speculative.

Sixth, continue efforts to study diet and cancer in populations with plausibly greater reliability in reporting as well as a wide range of dietary intake—it is the *ratio* of interindividual variation to intraindividual measurement error that determines the magnitude of relative risk distortion.

Seventh, continue to explore, as Hung et al. do, interactions between dietary and other potential cancer risk factors. Modest (but important) nutrition–cancer relationships may be obscured by associations with risk factors such as physical activity, use of nonsteroidal anti-inflammatory drugs, vitamins, or exogenous hormones, and genetic polymorphisms (15). Nutritional links may emerge only when interactions or risk factor stratum–specific results are addressed. Of course, such interaction work, now possible with the existence of large epidemiologic studies, is complicated by measurement error in the relevant variables.

We agree with Hung et al. that the prospective epidemiologic evidence to date does not provide strong support for a protective association between fruit and vegetable intake and cancer [although it is important to be alert to the possibility that findings

emerging from new, large cohort studies could shift the preponderance of the evidence, as may be occurring with the dietary fiber–colorectal cancer association (16,17,18)]. Our main point is that measurement error imbues nutritional epidemiologic findings with considerable uncertainty, which is only compounded with multivariate modeling. In other words, the evidence is simply inadequate at this time to determine whether fruit and vegetable intake confers modest protection against cancer. Researchers should recognize this uncertainty in nutrition and cancer epidemiology and do what it takes to move ahead, especially when it comes to improving exposure assessment in observational studies.

REFERENCES

- (1) Hung HC, Joshupura KJ, Jiang R, Hu FB, Hunter D, Smith-Warner SA, et al. Fruit and vegetable intake and the risk of major chronic disease. *J Natl Cancer Inst* 2004;96:1577–84.
- (2) International Agency for Research on Cancer. Fruits and vegetables. IARC handbooks of cancer prevention. Vol 8. Lyon (France): IARC Press; 2003.
- (3) Willett WC. Nutritional epidemiology. New York (NY): Oxford University Press; 1998.
- (4) Kipnis V, Midthune D, Freedman LS, Bingham S, Schatzkin A, Carroll RJ. Empirical evidence of correlated biases in dietary assessment instruments and its implications. *Am J Epidemiol* 2001;153:394–403.
- (5) Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano R, et al. The structure of dietary measurement error: results of the OPEN biomarker study. *Am J Epidemiol* 2003;158:14–21.
- (6) Carroll RJ, Ruppert D, Stefanski LA. Measurement error in nonlinear models. London (UK): Chapman & Hall; 1995.
- (7) Kipnis V, Freedman LS, Brown CC, Hartman A, Schatzkin A, Wacholder S. The effect of measurement error on energy adjustment models in nutritional epidemiology. *Am J Epidemiol* 1997; 146:842–55.
- (8) Frazier AL, Li L, Cho E, Willett WC, Colditz GA. Adolescent diet and risk of breast cancer. *Cancer Causes Control* 2004;15:73–82.
- (9) Bingham SA, Luben R, Welch A, Wareham N, Khaw KT, Day N. Are imprecise methods obscuring a relation between fat and breast cancer? *Lancet* 2003;362:212–4.
- (10) Schatzkin A, Kipnis V, Carroll RJ, Midthune D, Subar AF, Bingham S, et al. A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker based OPEN study. *Int J Epidemiol* 2003;32:1054–62.
- (11) Prentice R, Willett W, Greenwald P, Alberts D, Bernstein L, Boyd N, et al. Nutrition, physical activity, and chronic disease prevention: research strategies and recommendations. *J Natl Cancer Inst* 2004;96:1276–87.
- (12) Fung T, Hu FB, Fuchs C, Giovannucci E, Hunter DJ, Stampfer MJ, et al. Major dietary patterns and the risk of colorectal cancer in women. *Arch Intern Med* 2003; 163:309–14.
- (13) Kant AK, Schatzkin A, Graubard BI, Schairer C. A prospective study of diet quality and mortality in women. *JAMA* 2000;283:2109–15.
- (14) Trichopoulos A, Costacou T, Bamia C, Trichopoulos D. Adherence to a Mediterranean diet and survival in a Greek population. *N Engl J Med* 2003;348:2599–608.
- (15) Loktionov A. Common gene polymorphisms and nutrition: emerging links with pathogenesis of multifactorial chronic diseases. *J Nutr Biochem* 2003;14:426–51.
- (16) Bingham SA, Day NE, Luben R, Ferrari P, Slimani N, Norat T, et al. Dietary fibre in food and protection against colorectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC): an observational study. *Lancet* 2003;361:1496–501.
- (17) Peters U, Sinha R, Chatterjee N, Subar A, Ziegler RG, Kulldorff M, et al. Dietary fibre and colorectal adenoma in a U.S. colorectal cancer early detection program. *Lancet* 2003;361:1491–5.
- (18) Papas MA, Giovannucci E, Platz EA. Fiber from fruit and colorectal neoplasia. *Cancer Epidemiol Biomarkers Prev* 2004;13:1267–70.